What is claimed is:

1   1.   A method comprising:
2   modeling an audio-visual observation of a subject
3   using a coupled Markov model to obtain an audio-visual
4   model;
5   modeling a portion of the subject using an embedded
6   Markov model to obtain a portion model; and
7   determining first and second likelihoods of
8   identification based on the audio-visual model and the
9   portion model.

1   2.   The method of claim 1, wherein modeling the
2   audio-visual observation comprises using a coupled hidden
3   Markov model.

1   3.   The method of claim 2, wherein the coupled hidden
2   Markov model comprises a two-channel model, each channel
3   having observation nodes coupled to backbone nodes via
4   mixture nodes.

1   4.   The method of claim 1, further comprising
2   combining the first and second likelihoods of
3   identification.

1   5.   The method of claim 4, further comprising
2 weighting the first and second likelihoods of
3 identification.

1   6.   The method of claim 1, wherein the portion of the
2 subject comprises a mouth portion.

1   7.   A method comprising:
2   recognizing a face of a subject from first entries in
3 a database;
4   recognizing audio-visual speech of the subject from
5 second entries in the database; and
6   identifying the subject based on recognizing the face
7 and recognizing the audio-visual speech.

1   8.   The method of claim 7, further comprising
2 providing the subject access to a restricted area after
3 identifying the subject.

1   9.   The method of claim 7, wherein recognizing the
2 face comprises modeling an image including the face using
3 an embedded hidden Markov model.

1   10. The method of claim 9, further comprising
2 obtaining observation vectors from a sampling window of the
3 image.

1   11.   The method of claim 10, wherein the observation
2 vectors comprise discrete cosine transform coefficients.

1   12.   The method of claim 7, wherein recognizing the
2 face comprises performing a Viterbi decoding algorithm.

1   13.   The method of claim 7, wherein recognizing the
2 audio-visual speech further comprises detecting and
3 tracking a mouth region using vector machine classifiers.

1   14.   The method of claim 7, wherein recognizing the
2 audio-visual speech comprises modeling an image and an
3 audio sample using a coupled hidden Markov model.

1   15.   The method of claim 7, further comprising
2 combining results of recognizing the face and recognizing
3 the audio-visual speech pattern according to a
4 predetermined weighting to identify the subject.

1   16.   A system comprising:
2   at least one capture device to capture audio-visual
3 information from a subject;
4   a first storage device coupled to the at least one
5 capture device to store code to enable the system to
6 recognize a face of the subject from first entries in a

7   database, recognize audio-visual speech of the subject from
8   second entries in the database, and identify the subject
9   based on the face and the audio-visual speech; and
10   a processor coupled to the first storage to execute
11   the code.

1   17. The system of claim 16, wherein the database is
2   stored in the first storage device.

1   18. The system of claim 17, further comprising code
2   that if executed enables the system to model an image
3   including the face using an embedded hidden Markov model.

1   19. The system of claim 16, further comprising code
2   that if executed enables the system to model an image and
3   an audio sample using a coupled hidden Markov model.

1   20. An article comprising a machine-readable storage
2   medium containing instructions that if executed enable a
3   system to:
4   recognize a face of a subject from first entries in a
5   database;
6   recognize audio-visual speech of the subject from
7   second entries in the database; and
8   identify the subject based on recognizing the face and
9   recognizing the audio-visual speech.

1   21. The article of claim 20, further comprising
2 instructions that if executed enable the system to provide
3 the subject access to a restricted area after the subject
4 is identified.

1   22. The article of claim 20, further comprising
2 instructions that if executed enable the system to model an
3 image including the face using an embedded hidden Markov
4 model.

1   23. The article of claim 20, further comprising
2 instructions that if executed enable the system to model an
3 image and an audio sample using a coupled hidden Markov
4 model.